

RECEIVED

MAR 06 2002

TECH CENTER 1600/2900



1627

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/424,482C

TIME: 13:08:30

#20

Input Set : N:\Crf3\02212002\I424482A.raw

Output Set: N:\CRF3\03052002\I424482C.raw

```

1 <110> APPLICANT: CHOO, YEN
2      KLUG, AARON
3      ISALAN, MARK
4 <120> TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDE LIBRARY
5 <130> FILE REFERENCE: 71278/264974/BET
C--> 6 <140> CURRENT APPLICATION NUMBER: US/09/424,482C
7 <141> CURRENT FILING DATE: 2000-06-29
8 <150> PRIOR APPLICATION NUMBER: 09/424,482
9 <151> PRIOR FILING DATE: 1999-11-23
10 <150> PRIOR APPLICATION NUMBER: PCT/GB98/01510
11 <151> PRIOR FILING DATE: 1998-05-26
12 <150> PRIOR APPLICATION NUMBER: GB 9710809.6
13 <151> PRIOR FILING DATE: 1997-05-23
14 <160> NUMBER OF SEQ ID NOS: 114
15 <170> SOFTWARE: PatentIn version 2.1
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 9
19 <212> TYPE: DNA
20 <213> ORGANISM: Artificial Sequence
21 <220> FEATURE:
22 <223> OTHER INFORMATION: Description of Artificial Sequence: LIB-A DNA sorting
23      sequence
24 <221> NAME/KEY: modified_base
25 <222> LOCATION: (2)
26 <223> OTHER INFORMATION: a, t, c or g
27 <221> NAME/KEY: modified_base
28 <222> LOCATION: (3)..(4)
29 <223> OTHER INFORMATION: variable nucleotide
30 <400> SEQUENCE: 1
W--> 31      gnnncggcg
33 <210> SEQ ID NO: 2
34 <211> LENGTH: 9
35 <212> TYPE: DNA
36 <213> ORGANISM: Artificial Sequence
37 <220> FEATURE:
38 <223> OTHER INFORMATION: Description of Artificial Sequence: LIB-B DNA sorting
39      sequence
40 <221> NAME/KEY: modified_base
41 <222> LOCATION: (3)..(4)
42 <223> OTHER INFORMATION: variable nucleotide
43 <400> SEQUENCE: 2
W--> 44      gcnnccggcg
46 <210> SEQ ID NO: 3

```

ENTERED

9

9

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/424,482C

TIME: 13:08:30

Input Set : N:\Crf3\02212002\I424482A.raw

Output Set: N:\CRF3\03052002\I424482C.raw

```

47 <211> LENGTH: 9
48 <212> TYPE: DNA
49 <213> ORGANISM: Artificial Sequence
50 <220> FEATURE:
51 <223> OTHER INFORMATION: Description of Artificial DNA: LIB 1/2 sorting
52     sequence
53 <221> NAME/KEY: modified_base
54 <222> LOCATION: (5)..(9)
55 <223> OTHER INFORMATION: Any given nucleotide
56 <400> SEQUENCE: 3
W--> 57     gcggnnnnnn
59 <210> SEQ ID NO: 4
60 <211> LENGTH: 31
61 <212> TYPE: PRT
62 <213> ORGANISM: Artificial Sequence
63 <220> FEATURE:
64 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger
65 <221> NAME/KEY: MOD_RES
66 <222> LOCATION: (1)..(2)
67 <223> OTHER INFORMATION: Any amino acid
68 <221> NAME/KEY: MOD_RES
69 <222> LOCATION: (4)..(8)
70 <223> OTHER INFORMATION: Any amino acid
71 <221> NAME/KEY: MOD_RES
72 <222> LOCATION: (10)..(23)
73 <223> OTHER INFORMATION: Any amino acid
74 <221> NAME/KEY: MOD_RES
75 <222> LOCATION: (25)..(30)
76 <223> OTHER INFORMATION: Any amino acid
77 <221> NAME/KEY: MOD_RES
78 <222> LOCATION: (31)
79 <223> OTHER INFORMATION: His or Cys
80 <223> OTHER INFORMATION: Positions 1-2 may vary in length from
81     0-2 residues
82 <223> OTHER INFORMATION: Positions 4-8 may vary in length from
83     1-5 residues
84 <223> OTHER INFORMATION: Positions 10-23 may vary in length from
85     9-14 residues
86 <223> OTHER INFORMATION: Positions 25-30 may vary in length from
87     3-6 residues
88 <400> SEQUENCE: 4
W--> 89     Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
90           1             5             10             15
W--> 91     Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa
92           20             25             30
94 <210> SEQ ID NO: 5
95 <211> LENGTH: 24
96 <212> TYPE: PRT
97 <213> ORGANISM: Artificial Sequence

```

9

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/424,482C

TIME: 13:08:30

Input Set : N:\Crf3\02212002\I424482A.raw

Output Set: N:\CRF3\03052002\I424482C.raw

98 <220> FEATURE:

99 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger

100 <221> NAME/KEY: MOD_RES

101 <222> LOCATION: (1)

102 <223> OTHER INFORMATION: Any amino acid

103 <221> NAME/KEY: MOD_RES

104 <222> LOCATION: (3)..(6)

105 <223> OTHER INFORMATION: Any amino acid

106 <221> NAME/KEY: MOD_RES

107 <222> LOCATION: (8)..(10)

108 <223> OTHER INFORMATION: Any amino acid

109 <221> NAME/KEY: MOD_RES

110 <222> LOCATION: (12)..(16)

111 <223> OTHER INFORMATION: Any amino acid

112 <221> NAME/KEY: MOD_RES

113 <222> LOCATION: (15)..(16)

114 <223> OTHER INFORMATION: Any amino acid

115 <221> NAME/KEY: MOD_RES

116 <222> LOCATION: (18)..(19)

117 <223> OTHER INFORMATION: Any amino acid

118 <221> NAME/KEY: MOD_RES

119 <222> LOCATION: (21)..(23)

120 <223> OTHER INFORMATION: Any amino acid

121 <223> OTHER INFORMATION: Positions 3-6 may vary in length from

122 2-4 residues

123 <223> OTHER INFORMATION: Positions 8-10 may vary in length from

124 2-3 residues

125 <400> SEQUENCE: 5

W--> 126 Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa
 127 1 5 10 15

W--> 128 Leu Xaa Xaa His Xaa Xaa Xaa His
 129 20

131 <210> SEQ ID NO: 6

132 <211> LENGTH: 4

133 <212> TYPE: PRT

134 <213> ORGANISM: Artificial Sequence

135 <220> FEATURE:

136 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker

137 <400> SEQUENCE: 6

138 Thr Gly Glu Lys

139 1

141 <210> SEQ ID NO: 7

142 <211> LENGTH: 5

143 <212> TYPE: PRT

144 <213> ORGANISM: Artificial Sequence

145 <220> FEATURE:

146 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker

147 <400> SEQUENCE: 7

148 Thr Gly Glu Lys Pro

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/424,482C

TIME: 13:08:30

Input Set : N:\Crf3\02212002\I424482A.raw

Output Set: N:\CRF3\03052002\I424482C.raw

```

149      1      5
151 <210> SEQ ID NO: 8
152 <211> LENGTH: 26
153 <212> TYPE: PRT
154 <213> ORGANISM: Artificial Sequence
155 <220> FEATURE:
156 <223> OTHER INFORMATION: Description of Artificial Sequence: Consensus
157      structure sequence
158 <400> SEQUENCE: 8
159      Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Lys Ser Asp
160      1      5      10      15
161      Leu Val Lys His Gln Arg Thr His Thr Gly
162      20      25
164 <210> SEQ ID NO: 9
165 <211> LENGTH: 29
166 <212> TYPE: PRT
167 <213> ORGANISM: Artificial Sequence
168 <220> FEATURE:
169 <223> OTHER INFORMATION: Description of Artificial Sequence: Consensus
170      structure sequence
171 <400> SEQUENCE: 9
172      Pro Tyr Lys Cys Ser Glu Cys Gly Lys Ala Phe Ser Gln Lys Ser Asn
173      1      5      10      15
174      Leu Thr Arg His Gln Arg Ile His Thr Gly Glu Lys Pro
175      20      25
177 <210> SEQ ID NO: 10
178 <211> LENGTH: 6
179 <212> TYPE: PRT
180 <213> ORGANISM: Artificial Sequence
181 <220> FEATURE:
182 <223> OTHER INFORMATION: Description of Artificial Sequence: Leader peptide
183 <400> SEQUENCE: 10
184      Met Ala Glu Glu Lys Pro
185      1      5
187 <210> SEQ ID NO: 11
188 <211> LENGTH: 9
189 <212> TYPE: DNA
190 <213> ORGANISM: Artificial Sequence
191 <220> FEATURE:
192 <223> OTHER INFORMATION: Description of Artificial DNA: LIB 2/3 DNA
193      sorting sequence
194 <221> NAME/KEY: modified_base
195 <222> LOCATION: (1)..(5)
196 <223> OTHER INFORMATION: Any given nucleotide
197 <400> SEQUENCE: 11
W--> 198      nnnnnggcg
199
200 <210> SEQ ID NO: 12
201 <211> LENGTH: 9
202 <212> TYPE: DNA

```

9

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/424,482C

TIME: 13:08:30

Input Set : N:\Crf3\02212002\I424482A.raw

Output Set: N:\CRF3\03052002\I424482C.raw

```

203 <213> ORGANISM: Artificial Sequence
204 <220> FEATURE:
205 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger -DNA
206     interaction sequence
207 <400> SEQUENCE: 12
208     cgccccacgc
210 <210> SEQ ID NO: 13
211 <211> LENGTH: 9
212 <212> TYPE: DNA
213 <213> ORGANISM: Artificial Sequence
214 <220> FEATURE:
215 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger-DNA
216     interaction sequence
217 <400> SEQUENCE: 13
218     acgcccacg
220 <210> SEQ ID NO: 14
221 <211> LENGTH: 9
222 <212> TYPE: DNA
223 <213> ORGANISM: Artificial Sequence
224 <220> FEATURE:
225 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger-DNA
226     interaction sequence
227 <400> SEQUENCE: 14
228     gcgtgggcg
230 <210> SEQ ID NO: 15
231 <211> LENGTH: 9
232 <212> TYPE: DNA
233 <213> ORGANISM: Artificial Sequence
234 <220> FEATURE:
235 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger-DNA
236     interaction library designed sequence
237 <221> NAME/KEY: modified_base
238 <222> LOCATION: (7)..(9)
239 <223> OTHER INFORMATION: Any given nucleotide
240 <400> SEQUENCE: 15
W--> 241     acgccgnnn
243 <210> SEQ ID NO: 16
244 <211> LENGTH: 36
245 <212> TYPE: PRT
246 <213> ORGANISM: Artificial Sequence
247 <220> FEATURE:
248 <223> OTHER INFORMATION: Description of Artificial Sequence: LIB-A and
249     LIB-B Zinc finger
250 <400> SEQUENCE: 16
251     Met Ala Glu Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg
252     1             5             10             15
253     Arg Phe Ser Arg Ser Asp Glu Leu Thr Arg His Ile Arg Ile His Thr
254     20             25             30
255     Gly Gln Lys Pro

```

VERIFICATION SUMMARY

DATE: 03/05/2002

PATENT APPLICATION: US/09/424,482C

TIME: 13:08:31

Input Set : N:\Crf3\02212002\I424482A.raw

Output Set: N:\CRF3\03052002\I424482C.raw

L:6 M:270 C: Current Application Number differs, Wrong Format

L:31 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:44 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:57 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:89 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:91 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:126 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:128 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:198 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:241 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:271 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:287 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:303 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19